An Adaptive Rate-Based Congestion Control with Weighted Fairness for Large Round Trip Time Wireless Access Networks

Ladan Khoshnevisan
Electrical Engineering Department,
Faculty of Engineering,
University of Tehran,
Tehran, Iran.
l.khoshnevisan@ece.ut.ac.ir

Farzad R. Salmasi
Electrical Engineering Department,
Faculty of Engineering,
University of Tehran,
Tehran, Iran.
Farzad_rs@ieee.org

Vahid Shah-Mansouri
Electrical Engineering Department,
Faculty of Engineering,
University of Tehran,
Tehran, Iran.
vahids@ece.ut.ac.ir

Abstract— **In contrast with wired networks the time varying capacity in a wireless network makes the queue management more complicated than that in the wired ones. Moreover, any bit error in the acknowledgement packet sounds as the packet loss which may be misinterpreted as the congestion occurrence. Therefore, the window size will be reduced seriously which results in the resource wasting. In addition, large delay which is an ordinary aspect in a wireless network causes instability. On the other hand in a TCP connection the sources with small RTT value allocate the bottleneck link unfairly. The main contribution of our study is to design an adaptive robust rate-based queue management (ARRQM) based on the gradient projection internal model control for a large-delay wireless network to avoid the congestion contemporary with weighted fairness, maximum utilization and robustness against packet error rate and fading phenomena. To the authors' knowledge this is the first time that a queue management is designed for a wireless network with the mentioned advantages, simultaneously. The simulation results obtained via NS2 and Simulink confirm the analytical consequences.**

Keywords- **adaptive robust rate-based, congestion control, gradient projection, internal model control, weighted fairness, wireless access network**

I. INTRODUCTION

Due to the incremental number of the sources in a wireless access network, the congestion control methodologies have attracted many of the researchers during recent years. Transmission Control Protocol (TCP) is commonly used in the transport layer to control the data flow, based on the time-out mechanism and acknowledgements in a communication network through which the congestion is controlled only after its occurrence [1 and 2]. Consequently, to reduce the packet loss and the end to end delay, it is mandatory to design an active queue management (AQM) protocol which is implemented in the router [3]. AQM tries to control the congestion by adjusting the packet drop probabilities properly, to prevent buffer overflow and emptiness in the bottleneck link.

In a communication network, usually some sources compete on a common link, called bottleneck link, for sending packets. Fairness means to allocate an appropriate share of the bottleneck link to each of the users [4]. TCP connections inherently make the sources with small delay to allocate the bottleneck link unfairly [5]. Therefore, the users which are far from the station or have large RTT value (which is common in a wireless network) send packets through extremely small sending rates. Many approaches have been designed for the congestion control in the networks using the average window size such as random early detection (RED) [6], Smith [7], straightforward AQM (SFAQM) [8], internal model control [9], H∞ [10 and 11]. However, none of them can achieve fairness. To this aim some authors considered the rate-based procedures. For example Aweya et al. in [12] considered the aggregate rate through which the fairness could not be obtained. The optimization protocols are also considered in this area such as in [13], but only the state of the output queue is controlled and there is no control on the rates and input queue. Quet et al. in [14] tuned the sending rates to some constant values which had also no fairness characteristic because fairness does not mean determining a fixed value to each source [15]. Robustness against variation in the number of the sources is studied in [16] but weighted fairness is not obtained. Unal et al. in [17] compared some PI procedures with the consideration of the utilization characteristic in a network, but fairness is not investigated.

Due to the acknowledgement checking in a TCP connection, any bit error in the acknowledgement packet will be inferred as the packet loss. Therefore, it is mistaken as the congestion. Moreover, as in a wireless access network the data is transmitted through the open air, high Packet Error Rate (PER) is occurred which is the average expected number of errors in the packets during a period of time. In the case of PER existence TCP implementation reduce the window size to its half value and consequently the throughput value will be decreased seriously [18]. On the other hand the fading phenomena makes the bottleneck link capacity in the wireless networks to be varying in contrast with the wired communications [19]. Therefore, it is mandatory for the procedure to be robust against such disturbances. It is obvious that the node which handles the wireless links is the bottleneck. So the network topology which is considered in this paper consists of a bottleneck link and N TCP sources with some wireless access links. To the mentioned contributions in the wireless access network, the congestion control is still a vital problem although there are investigations of AQM. Some authors have proposed methods for the wireless networks in

[20] and [21] with time-varying capacity, but their procedure is not robust against the variation in the number of the sources and PER. Procedures proposed in [18] and [22] have taken into account PER in the wireless networks, but they have considered neither any bit error in the acknowledgement packet nor fading phenomena nor any variations in the number of the sources. Some papers such as [23] have considered wireless access link using average window size through which fairness has not been obtained.

The principal contributions of our study is listed as below.

- An Adaptive Robust Rate-based Queue Management (ARRQM) for a wireless access network is designed based on two degree of freedom internal model control (TDF-IMC) procedure.
- The proportion of the bottleneck link which is used by each source is considered as the utilization factors. The utilization factors are updated through an adaptive law obtained by the gradient method (**weighted fairness**).
- It is desired in a communication network that the sources use the maximum size of the bottleneck link bandwidth in order to avoid the link emptiness and overflow. Therefore, the projection method is accompanied with the gradient protocol through which both weighted fairness and **maximum utilization** are obtained.
- The advantages of the proposed procedure are weighted fairness, maximum utilization and robustness against PER and capacity variations caused by fading phenomena. Also the number of active sources and the utilization factors are determined adaptively.
- The procedure can tolerate large RTT effects [24] which is usual in the wireless networks.
- The analytical and simulation results verify these consequences and also they illustrate that the response has no oscillations and no delay jitter.
- To the authors' knowledge this is the first time that the fading problem, PER and variations in the number of the sources are considered in a large-delay wireless access network, simultaneously, and an adaptive robust rate-based congestion control is designed based on the gradient projection method with all of the mentioned advantages altogether.

The rest of the paper is organized as follows: The TCP/AQM dynamics are represented as a linearized fluid-flow model in Section II which is described as a multi-input single-output system. The adaptive two degree of freedom (ATDF) internal model (IM) controller is designed in Section II.B. The weighted fairness and maximum utilization are obtained via the gradient projection method in Section III. Some performance aspects such as the tracking problem and the robustness against disturbances and uncertainties are discussed in Section IV. Section V represents the simulation results obtained through NS2 and Simulink. Finally the conclusion remarks are denoted in Section VI.

II. SYSTEM MODEL

In a TCP implementation, the users with small RTT values allocate the bottleneck link to themselves. As a result it is mandatory to determine the link allocation for the users by controlling the sending rates to the specific values (**weighted fairness**). This aim can be obtained via the rate-based model. On the other hand, the procedure should be usable in the large RTT networks which is common in the wireless access links. On the other hand, occurring any error even in one bit of the acknowledgement packet is mistaken as the congestion event. To overcome such an obstacle, the packet error rate (PER) is modeled as the input disturbance which is formulated as below:

$$P_{PER} = 1-(1-P_{BER})^{S} \ , \tag{1}$$

where P_{BER} denotes the bit error rate and S represents the acknowledgement packet size.

Another important concept in a wireless network is the fading effect which causes the time varying capacity and packet loss. The resultant varying capacity is represented by the log-Normal distribution as below [25]:

$$C = M * \log\left(1+SNR|h|^2\right);\ f(y)=\frac{1}{2\,\sigma^2}e^{-\frac{y}{\sigma^2}}\ ;\ y=|h|^2 \ , \tag{2}$$

where SNR is the ratio of the average received signal to the noise power (dB), h denotes the time-varying channel gain, σ is the variance of y and the modulation gain is represented by M. In a wireless access network, the router which is responsible for the packet transmission through the wireless links is considered as the bottleneck node.

A. Fluid flow model of a TCP/AQM

Kunniyur and Srikant in [26] and [27] proposed a rate-based fluid flow model for a single bottleneck TCP source as below through which the slow start and timeout mechanism are ignored. The queue dynamic of the bottleneck link is proposed in [14] as represented in (4):

$$\dot{r}_i(t) = \frac{1}{d_i^2(t)} - \theta\, r_i(t)\, r_i(t-d_i(t))\, P_i(t-d_i(t)) \quad ;\ i=1,\dots,N\ , \tag{3}$$

$$\dot{q}(t) = -C(t) + \sum_{i=1}^{N} r_i(t) \quad ;\ i=1,\dots,N\ , \tag{4}$$

where $r_i(t)$ is the i'th TCP source rate (packets/seconds), $d_i(t)=q(t)/C(t)+T_{pi}$ is the i'th TCP source round trip time (RTT) (seconds), $P_i(t)$ is the TCP packet marking probability, θ is the parameter characteristic which demonstrates the type of the TCP source[1], $q(t)$ is the queue length, $C(t)$ is the time-varying link capacity which is caused by the fading phenomena (packets/sec) and N is the number of the active TCP sources. Since at every RTT, in congestion avoidance mode, the TCP window size increases by one and the sending rate is proportional to the window size, it is approximated that $r_i(t)=r_i(t-d_{0i}(t))$.

The steady state throughput, r_{0i}, is obtained via the steady state values of the packet marking probability (P_{0i}) and round trip time (d_{0i}) from (3) as below. The equilibrium point is:

[1] For instance θ=2/3 is specified for Reno TCP source.

$$r_{0i} = \frac{1}{d_{0i}}\sqrt{\frac{1}{\theta P_{0i}}}\ , \ q_0 = C_0(d_{0i} - T_{pi})\ , \tag{5}$$

Partitioning of the queue with special weights related to each users is called weighted fair queuing (WFQ) [28]. To obtain WFQ, some utilization factors (ρ_i , $i=1,..,N$) are designated to each source. Therefore, the portion of the bottleneck link which is considered for each source is determined by (6)

$$r_{0i} = \rho_i C_0\ . \tag{6}$$

By small signal linearization around the operating point [29], the following equation is obtained from (3) and (4):

$$\delta \dot{r}_i(t) = -2\theta r_{0i} P_{0i}\,\delta r_i(t) - \theta r_{0i}^2\,\delta P_i(t-d_0)\ \ ;\ i = 1,...,N\ , \tag{7}$$

$$\delta \dot{q}(t) = -\sum_{i=1}^{N}\frac{\rho_i}{d_{0i}}\delta q(t) + \sum_{i=1}^{N}\delta r_i(t) - \delta C(t)\ , \tag{8}$$

where δr_i , δP_i, δq and δC are the perturbed values around the operating point. The parameters with zero subscriptions represent the equilibrium point. In order to obtain the maximum bottleneck utilization, it is assumed that

$$\sum_{i=1}^{N}\rho_i = 1\ . \tag{9}$$

As a result, by taking Laplace transform form (7) and (8) the transfer function from the packet marking probability to the sending rate of the i'th TCP source (G_{tcp_i}) and from the sending aggregate rate to the queue length (G_{queue}) is obtained as below:

$$G_{tcp_i} = \frac{-\theta r_{0i}^2 e^{-d_{0i}s}}{s + 2\theta r_{0i} p_{0i}}\ ,\ G_{queue} = \frac{1}{s + \sum_{i=1}^{N}\frac{\rho_i}{d_{0i}}} \tag{10}$$

The schematic of the plant is considered as follows:

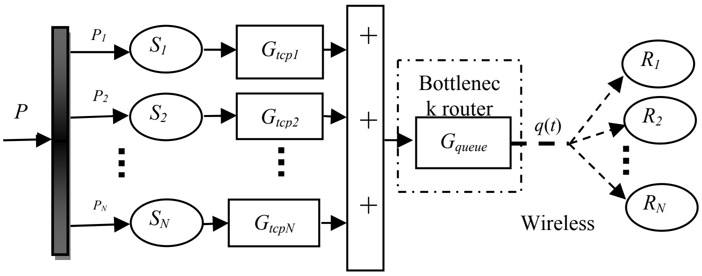

Figure 1 : schematic of the TCP network

where S_i denotes the i'th active source, R_i is the i'th receiver, P_i represents the drop probability of the i'th TCP packets and $q(t)$ is the queue length at time t. As it is illustrated in [7] the sources' sending rates are decoupled from each other. Therefore the system can be represented as a multi-input single-output system through the following equation.

$$\begin{aligned} &\begin{bmatrix}\delta r_1\\ ...\\ \delta r_N\end{bmatrix} = \begin{bmatrix}G_{tcp1} & 0 & 0\\ ... & ... & ...\\ 0 & 0 & G_{tcpN}\end{bmatrix}\begin{bmatrix}\delta P_1\\ ...\\ \delta P_N\end{bmatrix}\\ &\delta q(t) = G_{queue}\sum_{i=1}^{N}\delta r_i = \begin{bmatrix}G_{p_1} & ... & G_{pN}\end{bmatrix}\begin{bmatrix}\delta P_1\\ ...\\ \delta P_N\end{bmatrix} = G_P\begin{bmatrix}\delta P_1\\ ...\\ \delta P_N\end{bmatrix}\ , \end{aligned} \tag{11}$$

where G_{P_i} is as follows:

$$G_{P_i} = G_{queue}\,G_{tcp_i} = \frac{1}{s + \sum_{i=1}^{N}\frac{\rho_i}{d_{0i}}} \times \frac{-\theta r_{0i}^2 e^{-d_{0i}s}}{s + 2\theta r_{0i}P_i}\ . \tag{12}$$

B. ATDF-IM Controller

The adaptive two degree of freedom internal model control (ATDF-IMC) loop is depicted in the Figure 2 [30]:

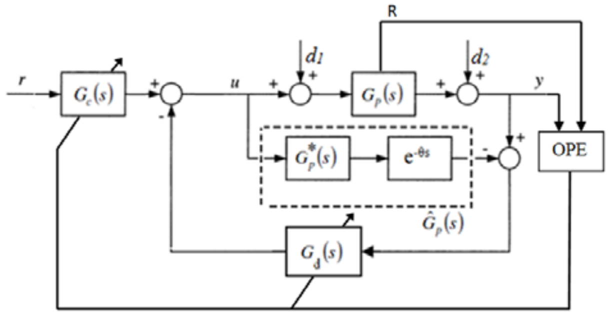

Figure 2 :TDF-IMC system,

where G_P is the multi-input single-output system to be controlled, r is the reference input, u is the packet drop probability, d_1 is the input disturbance which is the *PER* of the acknowledgement packets represented in (1), d_2 is the output disturbance which is due to the fading effects depicted in (2), R is the aggregate rate and y is the queue length which are used in the online parameter estimator (OPE) subsystem to estimate the utilization factors (ρ_i , $i=1,\ldots,N$). $G^*{}_P = \begin{bmatrix}G^*{}_{P_1} & ... & G^*{}_{P_N}\end{bmatrix}$ is the inner model of the system which is decomposed as below:

$$G^*{}_{P_i} = G^*{}_{M_i}\,G^*{}_{A_i}\ , \tag{13}$$

where $G^*{}_{M_i}$ is the minimum phase part and $G^*{}_{A_i}$ is the all pass part of the inner model which are represented in the following equation:

$$G^*{}_{M_i} = \frac{-\theta r_{0i}^2}{s + 2\theta r_{0i}P_i} \times \frac{1}{s + \sum_{i=1}^{N}\frac{\rho_i}{d_{0i}}};\ \ G^*{}_{A_i} = e^{-d_{0i}s}\ ; i = 1,...,N, \tag{14}$$

G_c is the set point tracking controller and G_d is the disturbance rejection controller which are designed adaptively based on the internal model control (IMC) procedure. It can be seen from Figure 2 that

$$\begin{aligned} u &= \left(1 + G_d(G_p - G_p^* e^{-\theta s})\right)^{-1} G_c\, r\\ &\quad + \left(1 + G_d(G_p - G_p^* e^{-\theta s})\right)^{-1}\left(-G_d\left(G_p d_1 + d_2\right)\right),\\ y &= \left(1 + G_d(G_p - G_p^* e^{-\theta s})\right)^{-1} G_p G_c\, r\\ &\quad + \left(1 + G_d(G_p - G_p^* e^{-\theta s})\right)^{-1}\left(1 - G_p^* e^{-\theta s} G_d\right)\left(G_p d_1 + d_2\right). \end{aligned} \tag{15}$$

Therefore, when the model matches the real process model i.e. $G_P = G_P^* e^{-\theta s}$, we have:

$$\begin{aligned} u &= G_c r - G_d(G_p d_1 + d_2),\\ y &= G_p G_c r + (1 - G_p G_d)(G_p d_1 + d_2). \end{aligned} \tag{16}$$

As a result, the design of the set point tracking controller is decoupled from the disturbance rejection controller which can be inferred from (17). The set point tracking controller is designed based on the IMC method as below:

$$G_c = \begin{bmatrix} G_{c\,1} \\ \dots \\ G_{c\,N} \end{bmatrix}, G_{c_i} = \hat{G}_{M_i}^{-1} \frac{\hat{\rho}_i}{(\lambda_C s + 1)^{n_i}} \; ; \; i = 1,\dots,N \, , \tag{17}$$

where $\hat{\rho}_i$ is the estimate of the utilization factors which will be obtained in the next section. The disturbance rejection controller is also designed based on the IMC procedure as

$$G_d = \begin{bmatrix} G_{d\,1} \\ \dots \\ G_{d\,N} \end{bmatrix}, G_{d_i} = \hat{G}_{M_i}^{-1} \frac{\hat{\rho}_i}{(\lambda_d s + 1)^{n_i}} \; ; \; i = 1,\dots,N \, , \tag{18}$$

where $\hat{G}_{M_i}$ is the minimum phase part of the inner model with the estimated value of the utilization factors, n_i is the relative degree of $\hat{G}_{M_i}$, λ_c and λ_d are the tuning parameters of the set point tracking controller and the disturbance rejection controller, respectively, which determine the speed of the response.

III. Adaptive Weighted Fair Queueing Procedure

In this section, the utilization factors ($\rho_i \; ; \; i=1,\dots,N$) are adapted through two adaptive laws. Weighted fairness is obtained through gradient procedure, through which a portion of the bottleneck link capacity is designated for each source. As the maximum utilization is an important factor for a communication network, this issue is assured by combining the projection method with gradient procedure.

A. WFQ

Basically, a TCP network causes the sources with short delays to fully utilize the bottleneck link with no fairness which prevents the other sources to use the link suitably. In a communication network, fair queueing (FQ) is proposed to the router which is designated to some queues, each of which are for a flow. To gain such a requirement, the proposed utilization factors ($\rho_i \; ; \; i=1,\dots,N$) in (6) are designed through gradient method in Theorem 1, adaptively, which are obtained in the OPE subsystem in Figure 2.

Theorem 1: For a TCP network with N sources and a bottleneck link, the following adaptive law for the utilization weight parameters causes the system to operate fairly.

$$\begin{bmatrix} \dot{\hat{\rho}}_1 \\ \dot{\hat{\rho}}_2 \\ \dots \\ \dot{\hat{\rho}}_N \end{bmatrix} = diag\begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_N \end{pmatrix} \times \begin{bmatrix} q^2/d_{01}^2 & q^2/d_{01}d_{02} & \dots & q^2/d_{01}d_{0N} \\ q^2/d_{02}d_{01} & q^2/d_{02}^2 & \dots & q^2/d_{02}d_{0N} \\ \dots & \dots & \dots & \dots \\ q^2/d_{0N}d_{01} & q^2/d_{0N}d_{02} & \dots & q^2/d_{0N}^2 \end{bmatrix} \begin{bmatrix} \hat{\rho}_1 - \rho_1 \\ \hat{\rho}_2 - \rho_2 \\ \dots \\ \hat{\rho}_N - \rho_N \end{bmatrix}, \tag{19}$$

where $diag(\gamma_1 \; , \; \dots \; , \; \gamma_N)$ is the weighting matrix which determines the convergence speed of the parameter estimation. The proof is provided in [30].

B. WFQ with maximum utilization

The ideal status of a communication network is that all the bottleneck link capacity to be used by the sources, which means maximum bottleneck utilization. As a result, it is mandatory to design the procedure based on the projection method contemporary with the gradient protocol through which both fairness and maximum utilization can be achieved. To this aim, (9) is the desired constraint to obtain the maximum utilization. Theorem 2 gives the required adaptive law.

Theorem 2: In a TCP network with N sources and a bottleneck link the following adaptive law for the utilization weight parameters causes the system to operate fairly with maximum utilization:

$$\begin{bmatrix} \dot{\hat{\rho}}_1 \\ \dot{\hat{\rho}}_2 \\ \dots \\ \dot{\hat{\rho}}_N \end{bmatrix} = \left(I - \frac{1}{\sum_{i=1}^{N} \gamma_i} \begin{bmatrix} \gamma_1 & \dots & \gamma_1 \\ \dots & \dots & \dots \\ \gamma_N & \dots & \gamma_N \end{bmatrix} \right) diag\begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_N \end{pmatrix} \times \begin{bmatrix} q^2/d_{01}^2 & q^2/d_{01}d_{02} & \dots & q^2/d_{01}d_{0N} \\ q^2/d_{02}d_{01} & q^2/d_{02}^2 & \dots & q^2/d_{02}d_{0N} \\ \dots & \dots & \dots & \dots \\ q^2/d_{0N}d_{01} & q^2/d_{0N}d_{02} & \dots & q^2/d_{0N}^2 \end{bmatrix} \begin{bmatrix} \hat{\rho}_1 - \rho_1 \\ \hat{\rho}_2 - \rho_2 \\ \dots \\ \hat{\rho}_N - \rho_N \end{bmatrix}, \tag{20}$$

where I is the identity matrix $\epsilon \; R^{N*N}$ and $\{\gamma_i \; ; \; i=1,\dots,N\}$ have been defined previously. The proof is provided in [30].

The reference value for the steady-state utilization weights are considered as below:

$$\rho_i^* = \frac{d_{0i}}{\sum_{j=1}^{N} d_{0j}} \quad ; i = 1,\dots,N \; , \tag{21}$$

where N is the number of active sources. The active sources will be obtained based on the following relationship:

$$\begin{aligned} &\text{If } r_i(t) > 0 \text{ then the i}^{\text{th}} \text{ source is active ;} \\ &\text{If } r_i(t) = 0 \text{ then the i}^{\text{th}} \text{ source is inactive .} \end{aligned} \tag{22}$$

Considering the summation of (21), (9) will be reached. Therefore, the weighted fair queuing with respect to the delay of each source contemporary with maximum utilization will be achieved via gradient projection procedure.

IV. System Performance

The major aim of our study is to design the active queue management to control the congestion by avoiding the buffer overflow in a large delay wireless network contemporary with the fading phenomena and PER in the acknowledgement packets. Also it is valuable for the procedure to establish the weighted fairness among the sources with different RTT values in spite of variations in the number of TCP sources. Some specific aspects such as tracking and robustness are scrutinized to investigate the proposed procedure.

A. Tracking

The important goal of rate-based congestion control in a wireless access network is to stabilize the steady-state queue length to obtain the congestion control together with the convergence of the steady-state flow rate to the desired value for the aim of the fair queuing. Therefore, the tracking problem is studied here. In this part the model is considered to be matched with the real plant. The input of the system is assumed as the step function with the desired value as below:

$$r_{ref} = \frac{q_0}{s} \tag{23}$$

In the disturbance free case it can be inferred from (16) that steady-state value of the output signal (y_∞) and the flow rate of each source are as follow via final theorem and certainty equivalence principle:

$$y_\infty = \lim_{s\to 0} s\left(\sum_{i=1}^{N} G_{pi}G_{ci}\right) r_{ref} = \lim_{s\to 0}\left(\sum_{i=1}^{N} \frac{\hat{\rho}_i\, q_0\, e^{-d_{0i}s}}{(\lambda_{Ci}s+1)^{n_i}}\right) = q_0 \tag{24}$$

$$r_{ss} = \lim_{s\to 0} s\, G_{tcp_i} G_C r_{ref} = \frac{\hat{\rho}_i q_{0i}}{\frac{q_{0i}}{C} + T_{pi}} \cong \hat{r}_{0i} \tag{25}$$

Conclusively, (24) illustrates that the congestion in the router will be avoided and **the system is stable for all RTT values.** Also it is concluded that the less λ_C (controller parameter) the more convergence speed. Moreover, it can be inferred from (25) that the weighted fairness is obtained through this procedure for a wireless network with large RTT value.

B. Robustness

It is desired for a queue management procedure in a wireless communication network, to be robust against PER in the acknowledgement packets (as the input disturbance), fading phenomena (considered as the output disturbance) and the parameter variations such as changing in the number of the sources. To analyze the robustness of the proposed scheme, nonzero disturbances are considered in Figure 2. Therefore, from equations (11), (12), (13), (16), (18) and by using the final value theorem, the steady state value of the output signal is obtained as follows:

$$\begin{aligned}
&\lim_{t\to\infty} y(t) = \lim_{s\to 0} sy(s) \\
&= \left(1 - G_p^*(0)\times G_d(0)\right)\times\left(G_p(0)\times\left(\lim_{s\to 0} sd_1(s)\right) + \lim_{s\to 0} sd_2(s)\right) \\
&= \left(1 - \sum_{i=1}^{N}\rho_i\right)\times\left(G_p(0)\times\left(\lim_{s\to 0} sd_1(s)\right) + \lim_{s\to 0} sd_2(s)\right) = 0
\end{aligned} \tag{26}$$

Consequently, the procedure rejects any kind of disturbance signal. On the other hand, in the disturbance free case, when the system parameters change, the real plant model does not match with the model $\hat{G}_p(s)$. Therefore, the output signal can be achieved via equations (11), (12), (15), (17), (18) and the final value theorem:

$$\begin{aligned}
&\lim_{t\to\infty} y(t) = \lim_{s\to 0} sy(s) \\
&= \left(1 + G_d(0)\times(G_p(0) - G_p^*(0))\right)^{-1}\left(G_p(0)\times G_c(0)\right)\times q_{ref} \\
&= \left(1 + \sum_{i=1}^{N}\left(\left(\frac{1}{K_{pi}}\right)\times\rho_i\times\left(\hat{K}_{pi} - K_{pi}\right)\right)\right)^{-1} \\
&\times\left(\sum_{i=1}^{N}\left(\hat{K}_{pi}\times\left(\frac{1}{K_{pi}}\right)\times\rho_i\right)\right)\times q_{ref} = q_{ref}\ ,
\end{aligned} \tag{27}$$

where $\hat{K}_{pi}$ is the real plant parameter which has been changed and K_{pi} is the model parameter as below:

$$\hat{K}_{pi} = \frac{(-\hat{r}_{0i})^3}{2p_{0i}}\times\frac{1}{\sum_{i=1}^{N}\frac{\rho_i}{\hat{d}_{0i}}},\ K_p = \frac{(-r_{0i})^3}{2p_{0i}}\times\frac{1}{\sum_{i=1}^{N}\frac{\rho_i}{d_{0i}}}. \tag{28}$$

The input signal is considered as in (23). The uncertain parameter is demonstrated by hat in (28). Therefore, the output signal converges to the desired value, despite there exists uncertainty in the plant parameters.

Consequently, the proposed procedure is robust against parameter uncertainties and disturbances such as fading and PER.

V. SIMULATION RESULTS:

In this section the simulation results of the proposed procedure in a wireless access network is illustrated. Two TCP sources are considered with different RTT values. The congestion avoidance and weighted fairness are scrutinized in a wireless network with PER and capacity variation. The network topology is considered as illustrated in Figure 1, with T_{p1}=1ms, T_{p2}=8ms, with the link bandwidth of 15Mb/s, and the packet size of 500 Bytes through which the bottleneck links capacity is 3750 packets/s. The bit error rate is distributed between $1.25*10^{-7}$ and $5*10^{-5}$ each of which is kept for 0.5s before jump to another and the acknowledgement packet size is assumed as 40 Bytes. The signal to noise ratio is 10dB, the time varying channel gain is considered as Gaussian(0,0.001) and the modulation gain is 700 [31]. The sending rates of each source is demonstrated in Figure 3 illustrates that the second source rate, which has higher RTT value, is more than the first one. **Accordingly, the weighted fairness is achieved through this procedure.**

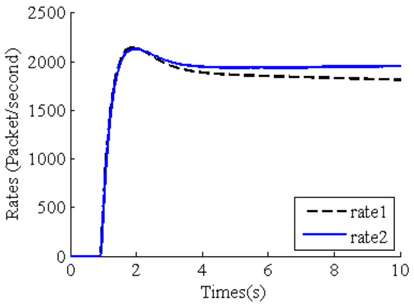

Figure 3 : The sending rates of each source.

To evaluate the designed procedure it is also simulated through Network Simulator 2 (NS2). To this aim it is required to discretize the controllers and the parameter estimation laws. Bilinear transformation is utilized to convert s to $\left(2(z-1)\right)/\left(T_s(z+1)\right)$, where T_s is the sampling time. Sampling frequency is considered around 10 times of the loop bandwidth [24]. The pseudo code which is used in the C++ source file can be generated via the difference equation obtained from the discrete time controllers. A wired-cum-wireless dumbbell network topology is considered for NS2 simulation with 50 FTP flows through which the delay of 25 of them are lower than the others. The desired value of the queue length is assumed as 125 packets. The bottleneck link capacity is considered as 1400 (pakcets/s), the packet size is 500 bytes and TCP Reno is assumed as the transport agent. Figure 4 and Figure 5 illustrate the queue length evolution when the RTT value of the bottleneck link is considered as 100 ms and 500 ms, respectively.

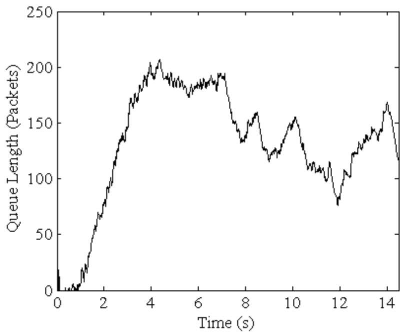

Figure 4 : The queue length evolution obtained via NS2 (RTT=100ms)

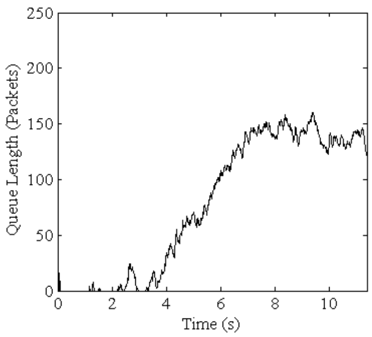

Figure 5 : The queue length evolution obtained via NS2 (RTT=500ms)

As it can be inferred from Figure 4 and Figure 5, the average queue length converges to the desired value in the networks with small and large RTT values. **Consequently, the congestion avoidance is obtained via this procedure**

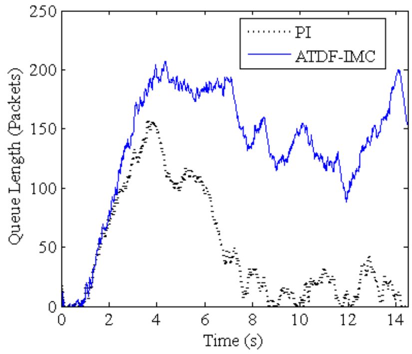

Figure 6: Comparison of queue length for ATDF-IMC (solid line) and PI (dotted line) procedures via NS2 (RTT=100ms).

It is illustrated in Figure 6 that the queue length does not converge to the desired value through PI procedure in a wireless access network. However, ATDF-IMC can lead to the desired value for the queue length evolution.

Therefore, simulation results, obtained from NS2 and Simulink, confirm that the procedure can avoid congestion in a wireless access network with the advantages of weighted fairness, maximum utilization and robustness against PER and fading phenomena. Moreover, the procedure can be proposed in the wireless networks with large RTT value effectively with no buffer overflow.

VI. CONCLUSION

Due to the increasing demand on the wireless networks and the capacity limitation, it is mandatory to design a congestion control. Opposed to the wired networks, the fading phenomena causes the capacity variations in the wireless networks. In addition, the packet error rate (PER) which is imposed by any bit error in the acknowledgement packet causes the packet loss. So this may be misunderstood as the congestion occurrence which enforces decreasing in the window size and wasting the system resources. Therefore, the queue management designing in a wireless network is more sophisticated than that in the wired ones. Moreover, large delay which is a common issue in the wireless networks makes the control system to be unstable. On the other hand the TCP implementation imposes the sources which has small delay to communicate with a large sending rate unfairly. Consequently, in this study an adaptive robust rate-based queue management (ARRQM) is designed based on the internal model control which can tolerate large RTT effects. The weighted fairness is achieved via the gradient method and the maximum utilization, which is an imperative characteristic in a communication network, is gained through the gradient projection protocol. It is proved through analytical results that the procedure is robust against any parameter uncertainties and external disturbances such as PER and capacity variations. As a result it can be used in the wireless access networks effectively. The procedure performance is also evaluated through NS2 and Simulink simulation results which approved the analytical proofs.

References

[1] R. Shorten, C. King, F. Wirth, D. Leith, Modelling TCP Congestion Control Dynamics in Drop-Tail Environments, Automatica, 43 (2007) 441-449.

[2] A. De Vendictis, A. Baiocchi, M. Bonacci, Analysis and Enhancement of TCP Vegas Congestion Control in a Mixed TCP Vegas and TCP Reno Network scenario, Perform. Eval, 53(2003) 225-253.

[3] V. Kushawaha, R. Gupta, Congestion Control for High-Speed Wired Networks: A Systematic Literature Review, J. Net. Comp. App, 45 (2014) 62-78. Doi: 10.1016/j.jnca.2014.07.005.

[4] C. So-In, R. Jain, A. K. Al Tamimi, Generalized Weighted Fairness and its Application for Resource Allocation in IEEE 802.16e Mobile WiMAX, Proc. 7th Annual IEEE Consumer Communications and Networking Conference, Las Vegas, NV, 1 (2010) 784-788.

[5] S. Floyd and V. Jacobson, Connection with multiple Congested Gateways in Packet-Switched Networks, Part 1: One-Way Traffic, ACM Comput. Commun. Rev., 21 (1991) 30-47.

[6] L. J. Pei, X. W. Mu, R. M. Wang, J. P. Yang, Dynamics of the Internet TCP-RED congestion control system, Nonlinear Analysis: Real World Applications, 12 (2011) 947-955.

[7] L. De Cicco, S. Mascolo, S. I. Niculesci, Robust Stability Analysis of Smith Predictor-Based Congestion Control Algorithms for Computer Networks, Automatica, 47 (2011) 1685-1692.

[8] Q. Xu, J. Sun, A Simple Active Queue Management Based On The Prediction Of The Packet Arrival Rate, J. Net. Comp. App., doi: 10.1016/j.jnca.2014.03.006, (2014).

[9] H. Wang, M. J. Cheng, Z. H. Tian, Robust Buffer Management Mechanism in Quality of Service Routers, J. Shanghai Jiaotong University, 16 (2011) 452-458.

[10] P. F. Quet, H. Ozbay, On the Design of AQM supporting TCP Flows Using Robust Control Theory, IEEE Trans. Autom. Control., 49 (2004) 1031-1036.

[11] K. Chavan, R. G. Kumar, M. N. Belur, A. Karandikar, Robust Active Queue Management for Wireless Networks, IEEE Trans. Control. Syst. Technol., 19 (2011) 1630-1638.

[12] J. Aweya, M. Ouellete, D. Y. Montuno, K. Felske, Design of Rate-based controllers for active queue management in TCP/IP networks, Comput. Commun., 31 (2008) 3344-3359.

[13] Lu, L., Du, H., Liu, R. P., CHOKeR: A novel AQM algorithm with proportional bandwidth allocation and TCP protection, IEEE Trans. Industr. Inform., 10, 637-644, (2014).

[14] P. F. Quet, B. Atashlar, A. Iftar, H. Ozbay, S. Kalyanaraman, Rate-Based Flow Controllers for Communication Networks in the Presence of Uncertain Time-Varying Multiple Time-Delay, Automatica, 38 (2002) 917-928.

[15] J. Mo, J. Walrand, Fair End-to-End Window-Based Congestion Control, IEEE/ACM Trans. Netw., 8 (2000) 556-567.

[16] Q. Xu, J. Sun, A Simple Active Queue Management based on the Prediction of the Packet Arrival Rate, J. Net. Comp. App., 42 (2014) 12-20. Doi: 10.1016/j.jnca.2014.03.006.

[17] H. U. Unal., D. Melchor-Aguilar, D. Ustebay, S. I. Niculescu, H. Ozbay., Comparison of PI Controllers Designed for the Delay Model of TCP/AQM Networks, Comp Commun, 36 (2013) 1225-1234.

[18] Wen, T., Chen, C., Ding, Z., Yang, T. C. (2008), 'A novel AQM Scheme for Wireless Networks with BER Estimation', Proc. 17th World Congress, IFAC, Seoul, Korea, 2919-2924.

[19] Tse, D., Viswanath, P., Fundamentals of Wireless Communication. Cambridge University Press, 2005.

[20] Chavan, K., Kumar, R. G., Belur, M. N., Karandikar, A., 'Robust Active Queue Management for Wireless Networks', IEEE Trans. Control Syst. Tech., 19, 1630-1638, 2011.

[21] Zheng, F., Nelson, J., 'An Hinf Approach to congestion control design for AQM Routers Supporting TCP Flows in Wireless Access Networks', Computer Networks, 51, 1684-1704, 2007.

[22] Wen, T., Ding, Z., 'Online Optimal Queue Scheme for Enhanced Transport over Wireless Links', 8th IEEE International Conference on Control and Automation, China, 1519-1524, 2010.

[23] L. Khoshnevisan, F. R. Salmasi, V. Shah-Mansouri, 'Robust queue management for TCP-based large round trip time networks with wireless access link', *IEEE Wireless Communications and Networking Conference (WCNC)*, 1309-1313, 2015.

[24] L. Khoshnevisan, F. R. Salmasi, A robust and high-performance queue management controller for large round trip time networks, Int. J. Sys. Sci, (2014) doi: 10.1080/00207721.2014.941959

[25] Tse, D., Viswanath, P., Fundamentals of Wireless Communication. Cambridge University Press, 2005.

[26] S. Kunniyur, R. Srikant, End-to-end congestion control: utility functions, random losses and ECN marks, in Proc. IEEE IFOCOM (2000) pp. 1323-1332.

[27] S. Kunniyur, R. Srikant, An adaptive virtual queue (AVQ) algorithm for active queue management, IEEE Trans. Netw., 12 (2004) 286-299.

[28] A. Kumar, D. Manjunath, J. Kuri, Communication Networking: An Analytical approach, Morgan Kaufmann, chapter 4, (2004) pp. 148-161.

[29] J. Aweya, M. Ouellette, D. Montuno, K. Felske, Rate-Based Proportional Integral Control Scheme for Active Queue Management, Int. J. Netw. Manag., John Wiley and Sons, 16 (2006) 203-231.

[30] L. Khoshnevisan, F. R. Salmasi, Adaptive rate-based congestion control with weighted fairness through multi-loop gradient projection internal model controller, IET Control Theory and Applications, (2015), doi: 10.1049/iet-cta.2015.0057.

[31] Chavan, K., Robust Queue Management over a Link with varying Capacity, Mater of Technology Degree Dissertation, Indian Institute of Technology Bombay, 2007.